Getting Started with Minitab

Minitab is a general-purpose statistical computing system designed especially for people with little or no previous computer experience. It is easy to use, flexible, and quite powerful. This handout describes how to access Minitab on the public UNIX workstation clusters at Duke and how to use Minitab to obtain descriptive statistics. Minitab is also available on some Duke PC and Macintosh computer clusters, and may be purchased at the Bookstore for home use.

To use any of the workstations you must first have an account on the Duke workstation network. Accounts have been created for all students who pre-registerred for STA110; others may obtain one at the Consulting Desk (room 136 North Building, 660-2983).

Find a free workstation at one of the UNIX clusters on campus:

• 133	Soc/Psych	(16 workstations)	• 139	Engineering	(10 workstations)
• 100	North	(13 workstations)	• 106	Teer Eng. Lib.	(30 workstations)
• 002CC	Bio/Sci	(6 workstations)	• 125	Carr	(6 workstations).

In addition, all the workstations are accessible from Duke terminals and PC clusters, and from home computers equipped with modems by dialing 684-4900 (8-n-1, 1200–9600bps). Read *Introduction to the DEC Workstations at Duke* ($2.50 at the Gothic Bookshop) or ask other users for more information.

To login to a workstation, jiggle the mouse and a "login" window should appear. Enter your `userid` and your `password` and hit `<RET>`. First, a *Message-of-the-day* window will appear. Read it and click the button on the bottom of this window to continue. When the outline of the *Session-Manager* window appears on the screen press the left mouse button, drag the window to the bottom of the screen, and then release the mouse button (you will need the session-manager window later to quit your session). Two more *Xterm* windows will appear on the screen. Using the mouse, position the cursor into one of these two windows. Now you are all set to start your session.

If this is your first usage of the system, you should immediately change your password. Do this by typing:

```
passwd
```

and following instructions. Be sure to remember your new password: you will need it to login next time.

Descriptive Statistics

One-Sample Data

As a first example, use Minitab to work problem 3.25 from the book. The data are in file `∼sta110/traffic.data`. First copy the data file to your account:

```
cp ∼sta110/traffic.data traffic.data
```

To make sure the `cp` command (think: *copy*) did what you intended it to do, use the command

```
ls
```

The `ls` command (think: *list*) shows a list of all files in your account— at this moment probably only the file `traffic.data` which you just copied. Now you are ready to start Minitab. Type:

```
minitab.
```

You should get the Minitab prompt `MTB >`.
First we need to read the data into Minitab:

```
set 'traffic.data' c1
```

The `set` command reads the data from file `traffic.data` into the Minitab column `c1`. Don't forget the single quotes around the file name!
Look at the column c1:

```
print c1
```

Get some descriptive statistics on the dataset:

```
describe c1
```

To understand the printout of the `describe` command, look at:

```
help describe
```

> *Question 1:* Using values of MIN, MAX, MEAN and MEDIAN, can you describe in words what the histogram for this data will look like?

Now type

```
hist c1
```

to plot a histogram for the traffic death data. Use `help hist` if you need some help to interpret the plot.

Another form of graphical summary for one-sample data is a box-plot:

```
boxplot c1
```

Boxplots display the main features of a data set. The middle half of the data set (25th to 75th percentiles) is represented by a box, and the median is marked with a "+". The extent of the data and the location of possibly extraordinary values are indicated on either side of the box with a "whisker". Type

```
help boxplot
```

for more help.

To leave Minitab use the command:

```
stop
```

k-Sample Data

Track etch devices [1] were under consideration as an inexpensive way to measure low-level radiation in homes. An experiment was done to study the usefulness of such devices. Measurements were taken on four different types of such devices: Filter, Membrane, Open Cup and Badge. Twenty devices of each type were put in a chamber and subjected to constant radiation. The amount of radiation measured by each device was recorded.

The data are given in the following table:

26	45	36	21
21	33	34	23
16	26	33	27
28	46	29	29
27	25	30	24
19	32	36	26
21	33	28	27
26	22	34	26
25	30	34	24
16	26	33	25
20	37	33	31
25	34	29	28
25	44	28	30
21	43	27	30
23	35	27	31
26	38	30	29
21	38	29	29
23	45	28	29
17	39	31	30
25	39	27	28

Before you start your MINITAB session, copy the data file into your account:

```
cp ~sta110/radon.data radon.data
```

Then start a MINITAB session

```
minitab
```

You can read the data into your MINITAB session by

```
read 'radon.data' c1 c2 c3 c4
names c1='FILTER' c2='MEMB' c3='CUP' c4='BADGE'
```

The first command reads the data file into the columns c1-c4. The second command assigns meaningful names to the data columns.

Question 2: Which type of data model does this experiment represent?

Question 3: How do the four devices compare?

The following MINITAB commands will be helpful in answering Question 3:

```
dotplot 'FILTER' 'MEMB' 'CUP' 'BADGE'
```

Again, to leave MINITAB, type

```
stop
```

Regression-Correlation Data

This is an exercise in using the correlation and regression facilities in Minitab. We will use actual data from a study of achievement, aptitude, and study habits. Copy the dataset into your account:

```
cp ~sta110/achieve.data achieve.data
```

Look at the data set by typing:

```
more achieve.data
```

The dataset has three columns and 30 rows. Each row corresponds to one subject and reports measurements

[1] See Ryan, Joiner and Ryan, p107 for more explainations.

for the variables ACHSCORE (column 1) – score on a standard achievement test; INTEL (column 2) – a measure of innate intellectual ability, and STUDHAB (column 3) – a measure of the quality of study habits:

```
47    20    55
46    19    40
44    10    56
46    23    81
36    20    49
59    22    76
50    23    49
58    21    63
52    17    69
57    17    88
31    14    46
41    24    48
49    20    31
66    30    63
49    17    57
34    14    50
57    21    61
44    14    26
53    21    51
40    18    25
53    16    64
51    17    24
57    24    69
54    26    66
46    21    52
38    12    40
70    25    107
53    19    41
43    15    35
56    23    46
```

If your *xterm* window is too short to display all rows, `more` will pause after filling the screen. Hit the space bar to continue. Start Minitab:

```
minitab
```

First, we need to read the data into Minitab columns:

```
read 'achieve.data' c1 c2 c3
```

This will read the columns of the data set into the Minitab columns `c1, c2` and `c3`. To assign names to the columns type:

```
names c1='ACHSCORE' c2='INTEL' c3='STUDHAB'
```

Don't forget the single quotes around the names!

Correlation Coefficient

Use the Minitab command `corr` to compute correlation coefficients:

```
corr 'ACHSCOR' 'INTEL' 'STUDHAB'
```

Again – don't forget the quotes around the variable names. The output of `corr` shows the correlation coefficients between each possible pair in one table. Type `help corr` for more details.

> *Question 4:* Discuss the correlation coefficients. What do they tell about the relationships between the variables.

> *Question 5:* What pattern do you expect for a scatterplot of ACHSCOR versus INTEL?

To verify your answer to the last question, type:

```
plot 'ACHSCOR' 'INTEL'
```

If you want to save the plot for later printing, use the `outfile` command and redo the plot:

```
outfile 'plot-score-intel.lis'
plot 'ACHSCOR' 'INTEL'
nooutfile
```

The `outfile` command tells Minitab to save all output – including the plot – in the file `plot-score-intel.lis` until the `nooutfile` directive.

Regression Line

To fit a regression line (i.e. "least squares line"), use the command:

```
regress 'ACHSCOR' 1 'INTEL'
```

The "1" tells Minitab that we are only using one "explanatory variable" (we will do regression with more than one explanatory variable later in the semester). For the moment, ignore most of the output except for the first two lines which state the fitted line (in the format $y = ax + b$) and the number labeled "R-sq" (which is simply the squared correlation coefficient).

> *Question 6:* For a student with intelligence score INTEL = 30, estimate his/her achievement score ACHSCOR using the estimated regression line.

Influential Points
What do you think will happen if we include an extra row in the data set on an individual with low ACHSCOR but high INTEL? Try:

```
insert 0 1 'ACHSCOR' 'INTEL'
```

The command tells Minitab that we want to insert an extra row of data between row 0 and 1 (i.e. before the first row of the data file). Minitab will respond with a prompt for the new data. Type the following numbers:

```
20 30
end
```

Then repeat the `corr` command:

```
corr 'ACHSCOR' 'INTEL'
```

Question 7: Why did the correlation coefficient decrease?

To find the new values for slope and intercept of the least-squares line:

```
regress 'ACHSCOR' 1 'INTEL'
```

Question 8: In what direction did the line change?

Using the Printer

You might want to print the plot which has been saved in the file `plot-achscor-intel.lis`. First make sure the file is there:

```
ls
```

To print the file, type:

```
lpr -Fl plot-achscor-intel.lis
```

The print should come out at the laser printer in the terminal room where you are sitting. If other users are sending output to the printer at the same time, the print jobs will be queued. You can look at the position of your job in the printer queue if you type:

```
lpq
```

To remove a printer job from the queue – if you have submitted the wrong file, or changed your mind – use:

```
lprm
```

Don't worry about accidentally removing other users' jobs – you can't, even if you wanted to.

Ending the Session

To log out of the machine, move the mouse cursor to the word **Session** in the *Session Manager* window, hold down the left mouse button, drag the pointer to the word **Quit**, and release the button. All your windows should disappear, and the login window should reappear on the screen. (Typing `exit` or `logout` will only log you out of an xterm window, but not from the machine!)